# On Computing Shannon's Sphere Packing Bound and Applications

M.Z. Ahmed[1], M.A. Ambroze and M. Tomlinson
Fixed and Mobile Communications Research,
University of Plymouth, Plymouth, PL4 8AA, United Kingdom
{mahmed,mambroze,mtomlinson}@plymouth.ac.uk

*Abstract* — **A new method to numerically evaluate Shannon's lower bound is presented in this paper. This new method is based on the Incomplete Beta function and permits the exact evaluation of the Sphere Packing Bound for a large range of code sizes, rates and probability of error. Comparisons with current standards (DVB–RCS, DVB–S2 and 3GPP) are also presented and discussed. It is shown that current standard coding schemes are about 0.6dB from the Shannon Limit corrected for Binary Signalling.**

## I. Introduction

The classic lower bound on Probability of codeword error was developed by Shannon[1]. This presented lower and upper bounds on the probability of codeword error for a spherical block code with equal energy, a certain average energy or an upper bounded value of energy per codeword. Of interest is the equal energy case and the lower bound on probability of codeword error. The exact value of the lower bound was explored and due to numerical instability and computation time only block lengths of a few hundreds were computed exactly (see [2, 3] and references therein). In this paper we use a new numerical method to compute this lower bound. We also present the lower bound for some standard coding schemes[4–6], and some improvements to these standards.

## II. Computing The Sphere Packing Bound

### A. Problem Formulation (as told by Shannon)

Shannon estimates the probability of error $P_e(M, n, \sqrt{P/N})$ for the best code of length $n$ containing $M$ codewords each of power $P$ and perturbed by noise of variance $N$. The number of codewords $M$ for the codes of interest is $2^k$ where $k$ is the number of information bits and $n$ is the number of codeword bits. The ratio $P/N$ is equal to $2RE_b/N_o$, where $R$ is code rate and $E_b/N_o$ is the energy per information bit and the factor 2 comes from the output of a matched receiver.

The sphere packing bound formulation has $M$ points over an $n$ dimensional sphere, with all points at a distance of $\sqrt{nP}$ from the origin. For decoding purposes, the $(n-1)$ hyperplanes which bisects the line connecting any two codewords forms a polyhedra (in fact pyramids) with the apex at the origin. The probability of error is the probability that the noise will move a point from the codeword location to an area outside the polyhedra. If the $i$th polyhedra has solid angle $\Omega_i$ ($\Omega_i$ is the area cut out by the pyramid on the $n$ dimensional spherical surface). The $i$th pyramid can be replaced by a cone centred at the origin with care taken such that the solid angles are the same. This deformation causes an increase in $P_e$ as moving the area to create a cone of the same solid angle results in moving smaller elements into the cone (increasing the $P_e$). A bound on the probability of error can then be found as

$$P_e \geq \frac{1}{M}\sum_{i=1}^{M} Q^*(\Omega_i) \tag{1}$$

where $\Omega_i$ is the solid angle of the $i$th pyramid and $Q^*(\Omega_i)$ is the probability of a point being carried outside a surrounding cone of solid angle $\Omega_i$. If $\Omega_0$ is the solid angle of an $n$ dimensional sphere, then

$$\sum_{i=1}^{M} \Omega_i = \Omega_0$$

The bound above can be further simplified by replacing $\Omega_i$ with the average $\frac{\Omega_0}{M}$ and thus

$$\begin{aligned} P_e &\geq \frac{1}{M}\sum_{i=1}^{M} Q^*(\Omega_i) = Q^*\left(\frac{\Omega_0}{M}\right) \\ \therefore \quad P_e &\geq Q^*\left(\frac{\Omega_0}{M}\right) \end{aligned} \tag{2}$$

This $P_e$ is Shannon's fundamental Lower Bound, and it is expressed in terms of half–cone angles $\theta$ rather than solid angles $\Omega$. Defining $Q(\theta)$ as the probability of being carried out of a half–cone angle $\theta$, and $\theta_1$ corresponds to $\Omega_0/M$; then

$$P_e \geq Q(\theta_1) \tag{3}$$

### B. Code Rate as a Function of the Cone Angle

The cone angle $\theta_1$ is such that the solid angle of the cone is $1/M = 2^{-k}$ times the full solid angle of a sphere. The solid angle of a cone with half angle $\theta$ can be computed as $\Omega(\theta)$

$$\Omega(\theta_1) = \frac{(n-1)\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \int_0^{\theta_1} (\sin\theta)^{n-2} d\theta \tag{4}$$

and $M$ times this should equal the surface area of an $n$ dimensional sphere of radius $r$ ($S_n(r)$)

$$S_n(r) = \frac{n\pi^{n/2} r^{n-1}}{\Gamma(n/2+1)} \tag{5}$$

Thus for a sphere of unit radius ($r = 1$), the ratio of the solid angle of the cone and the solid angle of the sphere is related to the number of messages as

$$\begin{aligned}\frac{1}{M} &= \frac{\Omega(\theta_1)}{S_n(r)} \\ &= \frac{(n-1)}{n}\frac{1}{\sqrt{\pi}}\frac{\Gamma(\frac{n}{2}+1)}{\Gamma\left(\frac{n+1}{2}\right)}\int_0^{\theta_1}(\sin\theta)^{n-2}d\theta\end{aligned}$$

For binary codes this results in

$$\int_0^{\theta_1}(\sin\theta)^{n-2}d\theta = \frac{\sqrt{\pi}}{2^k}\frac{n}{(n-1)}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma(\frac{n}{2}+1)}$$

Now the integral above can be computed using the relationships[7] page 149.

$$\begin{aligned}\int \sin^{2p} x \quad dx &= \frac{1}{2^{2p}}\binom{2p}{p}x + \\ &\quad \frac{(-1)^p}{2^{2p-1}}\sum_{k=0}^{p-1}(-1)^k\binom{2p}{k}\times \\ &\quad \frac{\sin(2p-2k)x}{2p-2k} \qquad (6)\end{aligned}$$

$$\begin{aligned}\int \sin^{2q+1} x \quad dx &= \frac{(-1)^{q+1}}{2^{2q}}\sum_{k=0}^{q}(-1)^k\binom{2q+1}{k}\times \\ &\quad \frac{\cos(2q+1-2k)x}{2q+1-2k} \qquad (7)\end{aligned}$$

where $p = n/2-1$ and $q = (n-3)/2$ and both $p$ and $q$ are integers.

These expressions can be used to evaluate the angle $\theta_1$ given an $(n,k)$ code.

### *C. Lower Bound on Probability of Codeword Error*

The lower bound $Q(\theta_1)$ is computed using a spherical Gaussian distribution which is equivalent to a noncentral $t$-distribution. The noncentral $t$-distribution is defined using $z$, $\delta$ and $x$ where $z$ and $x$ are Gaussian $(N(0,1))$ and $\delta$ is a constant. The distribution states that the ratio of $(z+\delta)$ to the rms of $f$ other random variable does not exceed $t$. Thus denoting this probability as $P(f,\delta,t)$ we have

$$P(f,\delta,t) = \Pr\left\{\frac{z+\delta}{\sqrt{\frac{1}{f}\sum_{i=1}^{f}(x_i)^2}} \le t\right\} \qquad (8)$$

This is a spherical Gaussian distribution with unit variance about a point $\delta$ from the origin in $f+1$ dimensional space. The probability $P(f,\delta,t)$ is the probability of being outside a cone from the origin having the line segment to the centre of the distribution as axis. Shannon showed that $f = n-1$, $\delta = \sqrt{n2RE_b/N_o}$ and $t = \sqrt{f}\cot(\theta) = \sqrt{n-1}\cot(\theta)$[1]. Thus we get the relationship between the probability of error and the parameters of the code and system as

$$Q(\theta) = P\left(n-1, \sqrt{n}\sqrt{\frac{P}{N}}, \sqrt{n-1}\cot(\theta)\right) \qquad (9)$$

The density function of the spherical Gaussian distribution is given in [8] as

$$\begin{aligned}f(t) &= \frac{f^{f/2}e^{-\delta^2/2}}{\sqrt{\pi}\Gamma(f/2)}\sum_{j=0}^{\infty}\left[\frac{\delta^j 2^{j/2}}{j!}\Gamma((f+j+1)/2)\times\right. \\ &\quad \left. t^j(f+t^2)^{-(f+j+1)/2}\right] \qquad (10)\end{aligned}$$

and using this can be shown that the cumulative density function of this can be evaluated using the incomplete beta function[9, 10].

$$\begin{aligned}P(f,\delta,t) &= 1-\frac{1}{2}e^{-\delta^2/2}\sum_{j=0}^{\infty}\left[\frac{(\delta/\sqrt{2})^j}{\Gamma\left(\frac{j}{2}+1\right)}\times\right. \\ &\quad \left. I_x\left(\frac{f}{2},\frac{j+1}{2}\right)\right] \qquad (11) \\ \text{with, } x &= f/(f+t^2) \qquad (12) \\ \text{and, } I_x(c,d) &= \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)}\int_0^x t^{c-1}(1-t)^{d-1}dt \qquad (13)\end{aligned}$$

(13) has positive terms and the computation of the probability of error becomes the evaluation of a truncated sum.

$$\begin{aligned}P(f,\delta,t) &= 1-\frac{1}{2}e^{-\delta^2/2}\sum_{j=0}^{N}\left[\frac{(\delta/\sqrt{2})^j}{\Gamma\left(\frac{j}{2}+1\right)}\times\right. \\ &\quad \left. I_x\left(\frac{f}{2},\frac{j+1}{2}\right)\right] + ERROR \qquad (14)\end{aligned}$$

The truncated sum results in an error which can be shown to be

$$|ERROR| \le \frac{1}{2}(1+\delta\sqrt{2})[1-Poi(N,2\delta^2)] \qquad (15)$$

where $Poi(N,2\delta^2)$ is a Poisson probability[10]. In this paper $N$ is large enough such that the computed probability of error computed to double precision accuracy is unchanged.

Thus the computed probability of error can be defined as

$$\begin{aligned}P(f,\delta,t) &= 1-\frac{1}{2}e^{-\delta^2/2}\sum_{j=0}^{N}\left[\frac{(\delta/\sqrt{2})^j}{\Gamma\left(\frac{j}{2}+1\right)}\times\right. \\ &\quad \left. I_x\left(\frac{f}{2},\frac{j+1}{2}\right)\right] \qquad (16)\end{aligned}$$

If we define the sum in (16) as

$$SUM = \sum_{j=0}^{2N+1} T_j G_j \qquad (17)$$

where

$$T_j = \frac{(\delta/\sqrt{2})^j}{\Gamma\left(\frac{j}{2}+1\right)}, \quad \text{and} \quad G_j = I_x\left(\frac{f}{2},\frac{j+1}{2}\right),$$

$$\text{with} \quad x = \frac{f}{f+t^2}$$

Also, defining

$$D_i = T_{2i}, \quad \text{and} \quad E_i = T_{2i+1}, \tag{18}$$

Then,

$$\begin{aligned} SUM &= \sum_{j=0}^{2N+1} T_j G_j \\ &= \sum_{i=0}^{N} T_{2i} G_{2i} + T_{2i+1} G_{2i+1} \\ &= \sum_{i=0}^{N} D_i G_{2i} + E_i G_{2i+1} \end{aligned} \tag{19}$$

All the terms of the summation above can be evaluated recursively. Defining $\lambda = \delta^2/2$

$$D_0 = 1, \quad E_0 = \delta\sqrt{2/\pi},$$

$$D_i = (\lambda/i)D_{i-1}, \quad E_i = (\lambda/(i+1/2))E_{i-1}$$

Defining

$$B(i) = I_x(b, a+i),$$

and applying the identity

$$I_x(d, c+1) = I_x(d,c) + C(c,d)x^d(1-x)^c,$$

with

$$C(c,d) = \frac{\Gamma(c+d)}{\Gamma(c+1)\Gamma(d)},$$

we get

$$\begin{aligned} B(i) &= I_x(b, a+i) \\ &= I_x(b, a+i-1) + C(a+i-1, b)x^b(1-x)^{a+i-1} \\ &= B(i-1) + S(i-1) \end{aligned}$$

where,

$$\begin{aligned} S(i) &= C(a+i,b)x^b(1-x)^{a+i} \\ &= \frac{\Gamma(a+b+i)}{\Gamma(a+i+1)\Gamma(b)} x^b(1-x)^{a+i} \\ &= (1-x)\frac{a+b+i-1}{a+i} S(i-1) \end{aligned}$$

By setting

$$\begin{aligned} B(i) &= G_{2i} = I_x(f/2, i+1/2), \quad \text{and} \\ BB(i) &= G_{2i+1} = I_x(f/2, i+1) \end{aligned}$$

(19) can then be written as

$$SUM = \sum_{i=0}^{N} [D_i B(i) + E_i BB(i)] \tag{20}$$

with all four terms computed recursively.

A Logarithmic version of this algorithm is given in table 1

This algorithm can be used to evaluate the exact sphere packing lower bound for $k$ of thousands and arbitrary rate codes.

1. Input
   $f = n-1$,
   $\delta = \sqrt{n}\sqrt{2RE_b/N_o}$,
   $t = \sqrt{n-1}\cot(\theta)$, and
   $N$
2. Evaluate
   $\lambda = \delta^2/2$, and
   $x = f/(f+t)$
3. Evaluate
   $B = \ln I_x(f/2, 1/2)$,
   $BB = \ln I_x(f/2, 1)$,
   $D = 0$,
   $E = \ln(\delta\sqrt{2/\pi})$,
   $S = \ln 2 + \ln\Gamma((f+1)/2) - \ln\Gamma(f/2) - \ln(\sqrt{\pi}) + (f/2)\ln(x) + (1/2)\ln(1-x)$,
   $SS = \ln\Gamma(1+f/2) - \ln\Gamma(f/2) + (f/2)\ln(x) + \ln(1-x)$, and
   $SUM = logsum(D+B, E+BB)$
   where $\ln I_x(a)$ is the Log Incomplete Beta Function
   $\ln\Gamma(a)$ is the Log Gamma function and
   $logsum(a,b) = \ln(e^a + e^b) = a + \ln(1+e^{b-a})$
   assuming $a > b$.
4. For each $i = 1, 2, \cdots, N$
   $B = logsum(B, S)$
   $BB = logsum(BB, SS)$
   $D = \ln(\lambda) + D - \ln(i)$
   $E = \ln(\lambda) + E - \ln(i+0.5)$
   $SUM = logsum(logsum(SUM, D+B), E+BB)$
   $S = \ln(1-x) + \ln(f+2i-1) - \ln(1+2i) + S$
   $SS = \ln(1-x) + \ln(f+2i) - \ln(2+2i) + SS$
5. Return probability of error $1 - e^{(-\lambda + SUM + \ln(0.5))}$

Tab. 1: Logarithmic Version of Algorithm to Compute (9)

### *D. Correction for Binary Transmission*

Wideband Binary Gaussian Channels are known to have a loss due to binary transmission[11]. This loss comes from equating the capacity of the two–input Gaussian channel to the results from information theory.

The capacity of a Gaussian Channel with $\alpha = \sqrt{2(RE_b)/N_o}$ in nats/symbol is given by

$$C(\alpha) = \alpha^2 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} \log\cosh(\alpha^2 + \alpha y) dy \tag{21}$$

and a bit error probability of $p$ is achievable if

$$\left(\frac{\log(2)}{\log(e)} R\right)(1 - H(p)) \le C(\alpha) \tag{22}$$

with $H(p) = -p\ln(p) - (1-p)\ln(1-p)$ the natural entropy function and $R$ in bits/symbol and the factor $\frac{\log(2)}{\log(e)}$ converts the bits/symbol to nats/symbol.

The capacity for non–binary signalling is given by

$$C_\infty = \frac{1}{2}\log_e(1 + 2(RE_b)/N_o) \tag{23}$$

and using this we can compute the $E_b/N_o$ required by non–binary signalling. The difference between these two values of $E_b/N_o$ gives the loss for binary signalling. The $E_b/N_o$ difference and its dependence on code rate for low error rates (below $10^{-4}$) is shown in the figure below.

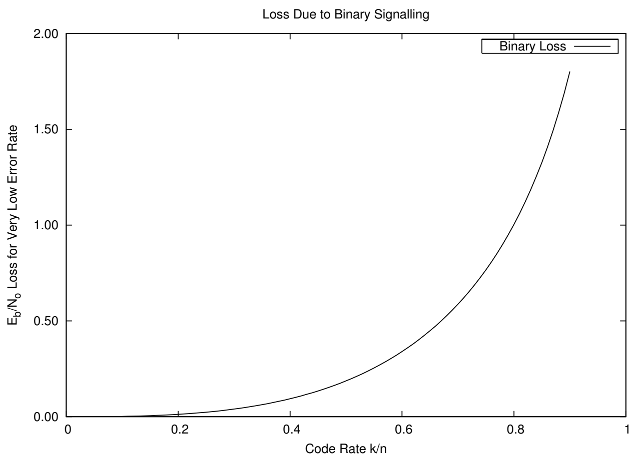


Fig. 1: Loss Due to Code Rate on Binary Channel

The sphere packing bound needs to be corrected for the loss due to binary signalling, to apply it to the simulations and the results that we have. This correction is done as an offset to $E_b/N_o$ as computed by the procedure above.

## III. Results

As an example, we have calculated the Shannon limit for the following standard codes: the MPG size $(3008, 1504, 19)$ DVB-RCS turbo code, the $(64800, 32400)$ DVB-S2 Hughes Network Systems LDPC and the $(1152, 384)$ 3GPP turbo code. The performance of the DVB-RCS code is shown in Figure 2, together with the Sphere Packing Bound corresponding to code sizes (3008,1504) and the correction for binary signalling. The performance of an improved turbo code $(3008, 1504, 25)$ using the same component codes and a code matched interleaver is also shown in this figure. It can be seen that the DVB-RCS turbo code is about 0.7 dB away from the Sphere Packing Bound and 0.5 dB away from the bound corrected for binary signalling for $P_e \geq 10^{-4}$. However, for $P_e < 10^{-4}$, the DVB-RCS code performance shows a pronounced error floor due to the relatively low minimum distance $d_{min} = 19$ and high multiplicity $a_{w=19} = 376$. This is due to the interleaver construction technique. We have designed a code matched interleaver for the same component codes and code size which results in a turbo code with $d_{min} = 25$ with multiplicity $a_{w=25} = 30$. This code has no error floor down to a probability of error $P_e = 10^{-6}$ and for this range of probabilities is 0.5 dB from the sphere packing bound corrected for binary signalling.

Figure3 shows the Sphere Packing Bound and the simulated performance of the $(1152, 384)$ 3GPP code. This code has a lower rate and as a consequence the binary signalling correction is smaller. Two curves are presented for this code: the basic iterative decoding curve and the performance with improved iterative decoding using the RVCM technique presented in [12–14]. It can be seen that this code is around 0.6 dB from the Sphere Packing Bound corrected for binary signalling.

The DVB-S2 code performance shown in Figure 4 is much longer block length than the previous examples. However, it is interesting to note that this code is also around 0.6 dB away from the corrected Sphere Packing Bound.

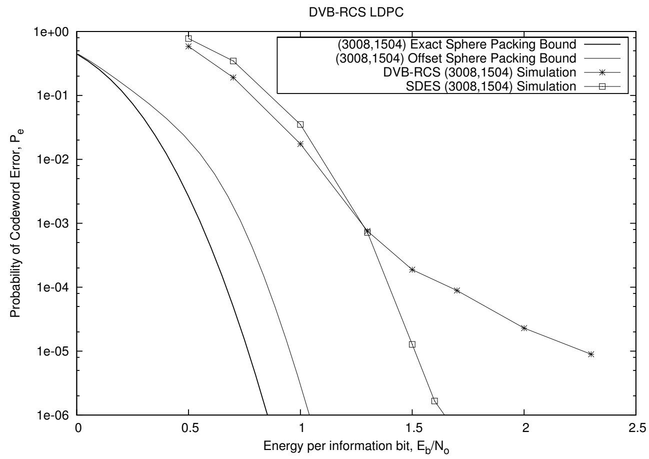


Fig. 2: DVB-RCS BPSK $(3008, 1504)$ LDPC

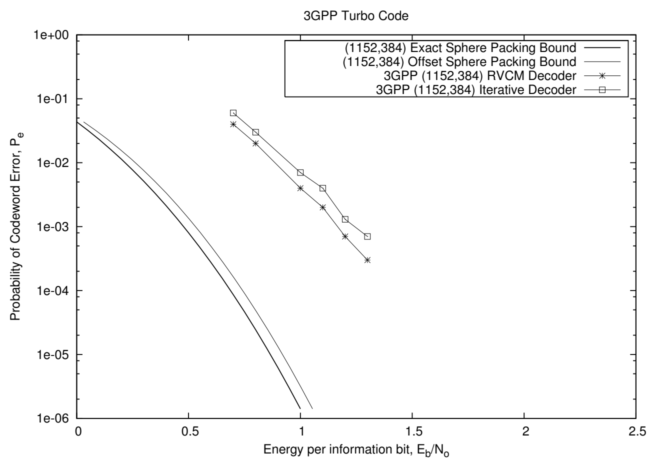


Fig. 3: 3GPP BPSK $(1152, 384)$ Turbo Code

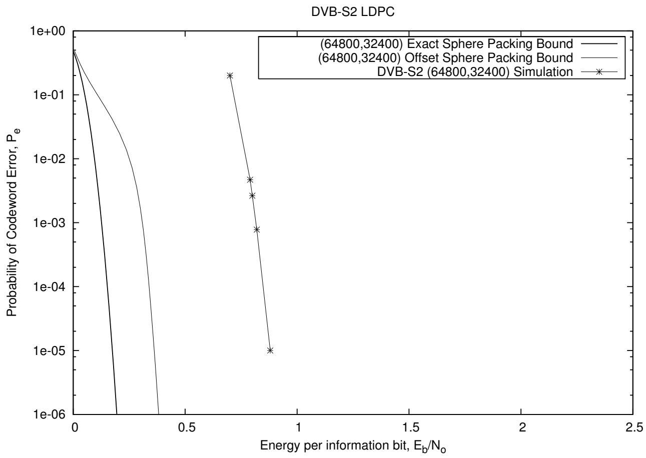


Fig. 4: DVB-S2 QPSK $(64800, 32400)$ LDPC

## IV. Conclusions

This paper has presented a new method for computing Shannon's Sphere Packing Bound based on using the Incomplete Beta Function. This new method enables the evaluation of the bound exactly for large values of $k$ and $n$. The calculation has been illustrated for different code rates and sizes corresponding to several standards such as DVB–RCS, DVB–S2 and 3GPP. We have noticed that regardless of code size and code rate the performance is about 0.6dB from the corrected Shannon Lower Bound at a codeword error rate of $10^{-6}$.

## References


[1] C. E. Shannon, "Probability of error for optimal codes in a gaussian channel," *The Bell System Technical Journal*, vol. 38, pp. 611–656, May 1959.

[2] S. Dolinar, D. Divsalar, and F. Pollara, "Code performance as a function of block size." TMO Progress Report 42-133, May 1998.

[3] M. Tomlinson, G. Wade, P. V. Eetvelt, and A. Ambroze, "Bounds for finite block–length codes." IEE Proc. Commun Vol 149, No 2, pp 65–69, Apr. 2002.

[4] ETSI EN 301 790 v1.4.1 (2005-04), *Digital Video Broadcasting (DVB); Interaction channel for satellite distribution systems.*

[5] ETSI EN 302 307, *Digital Video Broadcasting (DVB); S2.*

[6] ETSI TS 151 010, *3GPP.*

[7] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products.* Academic Press, 6th ed., July 2000. ISBN 0 1229 4757 6.

[8] Johnson and Kotz, *Continuous Univariate Distributions*, vol. 2. John Wiley & Sons, Inc.

[9] R. V. Lenth, "Cumulative distribution function of the non–central $t$ distribution," *Journal of Applied Statistics*, vol. 38, no. 1, pp. 185–189, 1989.

[10] H. O. Posten, "A new algorithm for the noncentral $t$ distribution function," *Journal of Statistical Computation and Simulation*, vol. 51, pp. 79–87, 1994.

[11] S. A. Butman and R. J. McEliece, "The ultimate limits of binary coding for a wideband gaussian channel," *JPL Deep Space Network Progress Report*, vol. 44, no. 22, pp. 78–80.

[12] E. Papagiannis, M. Ambroze, and M. Tomlinson, "Analysis of non convergent blocks at low and moderate SNR in SCCC turbo schemes," *SPSC 2003, $8^{th}$ Intl Workshop on Sig Proc for Space Comm, Catania, Italy*, pp. 121–128, Sept. 2003.

[13] E. Papagiannis, M. Ambroze, and M. Tomlinson, "Approaching the ML performance with iterative decoding," *IEEE IZS, Zurich*, pp. 220–223, Feb. 2004.

[14] N. Varnica and M. Fossorier, "Belief propagation with information correction : Imporved near maximum–likelihood decoding of low–density parity–check codes," *ISIT, Chicago, USA*, p. 343, July 2004.